多维度视角下百度百科词条评价指标构建*

何振宇 董祥祥 朱庆华
南京大学信息管理学院 南京 210023

摘要:[目的/意义]将用户使用行为纳入百科词条评价体系中,从多维度视角对百度百科词条进行综合评价。[方法/过程]结合国内外学者的研究成果,基于 4 个维度,选取了 12 项量化指标作为研究指标,对百度百科词条作了综合评价。[结果/结论]用户使用行为指标与链接指标在词条综合评价中起到相对重要的作用;研究丰富了百科词条研究的评价体系,但局限是未考虑丰富度、严谨性等难以准确量化的特征。

关键词:百度百科词条 灰色关联分析 熵权法 综合评价
分类号:G250
DOI:10.13266/j.issn.0252-3116.2019.12.015

1 引言

随着 Web2.0 技术的应用,网络百科作为典型的 Wiki 类产品,取得了迅猛的发展,其中百度百科正是国内网络百科中最具代表性的产品。词条(或条目)是网络百科的最小单位,截至 2019 年 2 月 8 日,百度百科已有 15 906 266 个词条,共有 670 万人参与了编写。词条是百度百科的核心内容,百度百科将词条细分为了艺术、科学、自然、文化、地理等 11 大类,词条结构非常固定化,包含词条名称、词条摘要、表格形式的词条基本信息、目录、词条正文、参考注释、词条标签等等。

网络百科内容主要由用户生成,受词条编辑者自身的主观性影响较大,导致网络百科用户使用者与管理者并不能很好地衡量词条的质量。虽然百度百科制定专有的审核制度,但词条信息质量很大程度上取决于广大编辑者的素质,由于编辑者在编辑过程中并不会完全遵循严谨的结构化过程,加上网络百科的词条数量如今呈现爆炸式增长的趋势,导致词条质量愈发参差不齐,也影响了百科用户的正常使用体验。百科词条的综合评价研究,从始至终是一个重要的问题。

2 研究现状

国内外目前对于网络百科的质量评价的研究主要集中在以下几方面:

(1)定性评价研究。主要是参考传统百科词条的标准,对要评价的词条作定性分析,利用专家或个人的知识进行人工评审。许多学者研究了影响词条质量的因素,创建了一套客观的评价标准。维基百科网站[1]自己就提出了详细的条目质量评级标准,通过同行评审的方式,将词条分为典范、甲、优良、乙、丙、初、小作品等不同的级别;B. Stvilia[2]等探索了不同文化和社区背景对信息质量评估的影响,以及将一些基于文章编辑的指标用于不同情境下的维基百科自动化质量测量的可行性;S. Lichtenstein[3]等建立了维基百科的知识生产模型并分析了模型的信息质量;O. Arazy[4]等建立了一个理论模型,解释了成员知识库的多样性、任务冲突及成员角色三个因素如何相互作用以确定维基百科词条的质量,并对 96 篇词条进行了实证研究;J. Liu[5]等分析了编辑者之间的合作和词条质量之间的关系,并且进行了实证研究,有助于改进维基百科的质量;E. Yaari[6]等选择 64 名用户评估维基百科的 5 篇词条质量,探讨了信息消费者在协作式的协作环境中对内容质量进行评估的方式;K. Osman[7]等通过扎根的方法,分析了 147 个关于维基百科质量的对话,探讨了词条冲突对维基百科质量的作用;D. Lewandowski[8]等总结了对百科词条质量的扩展讨论,开发了一种评估维基百科词条的启发式方法,并且将维基百科网页在搜索

* 本文系国家自然科学基金面上项目"协同视角下社会化搜索的形成机制与实现模式研究"(项目编号:71473114)研究成果之一。
作者简介:何振宇(ORCID:0000-0002-6002-7884),博士研究生,通讯作者,E-mail:demandsupply@163.com;董祥祥(ORCID:0000-0003-0273-3817),硕士研究生;朱庆华(ORCID:0000-0002-4879-399X),副院长,教授,博士,博士生导师。
收稿日期:2018-10-31 修回日期:2018-12-04 本文起止页码:114-120 本文责任编辑:杜杏叶

引擎的测试结果,与实验人员的人工判断相关性的结果进行了比较分析。丁敬达[9]通过三个方面具体考察了维基百科词条的信息质量内涵,总结了国内外对于维基百科词条信息质量的研究现状和不同的评价指标,并且构建了一种包含概念、关系、分类和方法学的多维结构的启发式评价框架,对维基百科词条进行信息质量评价;赵文轩[10]则详细分析了网络百科的特点优势、工作模式、资源对比三个方面,探讨了词条的信息质量,给出了优质词条的质量评价标准,并详细阐述了影响网络百科词条质量的六大因素;黄令贺[11]研究了网络百科信息质量的动态演进特点,提出了针对性的评价策略,注重过程评价、考虑生命周期特性、需要第三方积极介入。

(2)评价指标体系的构建。这类研究基于统计指标,通过量化的评价方法,建立了针对词条质量的详细具体的统计指标体系。A. Lih[12]根据已建立的百科全书分类法,从编辑历史信息的角度,提取出了"多样性"和"严谨性"两个评价指标,用以评判词条质量;F. Chevalier[13]等通过可视化的方法,提出了五个维基百科文章的成熟度和质量评估指标,保持着质量准确性的同时显著缩短了评估时间;金燕、周婷[14]等研究了普遍意义上的协同内容创建系统,采用层次分析法建立了具体的评价指标体系,从百度百科的实际情况出发,构建了针对性的百度百科的评价指标体系,并进行了验证;张博[15]等采用层次分析法,综合了用户需求与内容特征,以维基百科为实例,建立了4个层面、14个指标的内容质量评估模型;金燕[16]从用户体验的角度,建立了协同内容创建系统的质量评价指标体系,并且在信息价值、系统性能、用户满意度3个方面提出了质量保证措施。

(3)自动评价研究。随着计算机技术的发展,自动评价方式也被引入到网络百科的词条评价上来。通过选取合适的词条特征,并结合其他属性,使用计算机机器学习算法,自动化完成词条的评级工作。M. Warnckewang[17]等在B. Stvilia[18]等人的工作基础上,最终筛选出5个特征,同样使用了决策树算法对词条质量进行分类;J. E. Blumenstock[19]提出一种简单的仅通过单词数量来测量维基百科词条质量的方法,并且通过实证检验,表明其效果要好过许多复杂的方法;H. Dalip[20]等探索了大量的质量指标,通过机器学习的方式,将评价指标统一到同一个框架中,并且发现了与质量评价最相关的指标和无关的指标;L. Calzada[21]等提供了一个简单的信息质量模型,主要研究了"稳定"和"有争议的"的维基百科文章,并与人工评估的信息质量进行对比。裘江南[22]等提出了15个可量化的客观指标,利用决策树C4.5分类算法,构建了一个用于评价维基百科页面信息质量的自动分类模型;李欣奕[23]选取内容特征和编辑历史特征作为评判词条质量的核心指标,使用基于SVM的分类方法和PageRank算法对优质与劣质词条做出区分和排序;仝召娟[24]等选取了三个网页特征值,通过自动化处理和提取网页特征,选择高质量的分类器,实现了网页质量的自动化评价;肖奎,李兵[25]等考虑了词条属性和编辑者的行为,筛选出15个词条属性与4个用户属性,通过分类和利用余弦相似度计算词条质量来确定维基百科词条的等级。

(4)基于模型的排序评价。基于特定的模型,自动地计算词条的质量值并排序,从而完成对词条质量的评价过程。M. Hu[26]等利用维基百科词条编辑历史中的词条及其贡献者之间的交互数据,提出了三种质量检测模型:Basic Model、PeerReview Model、ProbReview Model,采用NDCG@k指标进行算法评价,取得了不错的表现。H. Zeng[27]、E. Lim[28]、P. Dondio[29]等分别使用不同的模型,从编辑行为的角度对词条质量进行计算,得到了词条质量的排序结果;T. Wohner[30]等提出了一种基于信息生命周期的方法,对维基百科词条质量进行评价;S. Javanmardi[31]等通过对维基百科词条的内容质量的动态演变进行建模,评估了维基百科的特色和非特色文章,并且通过CalSWIM做了案例分析;Y. Suzuki[32]从词条内容的存活率角度对词条质量进行评价,并开发出了一套词条排序方法;X. Li[33]等通过试用文章和编辑者之间的关系,开发了几种模型对维基百科文章进行排名,并且提出使用手动评估来辅助自动评估是可行的质量评估解决方案。

有些学者还利用链接分析的思想,可以计算词条质量的排序。D. Wilkinson等[34]等利用PageRank算法思想先对词条进行分级,然后分析同一等级中的词条质量,进而发掘优质词条的共性;肖奎,罗保山[35]等通过限定领域,应用HITS算法的思想,利用特定领域内编辑者的信誉度,自动计算后得出词条的质量值。K. Wu[36]等通过对六种不同语言维基百科的修订历史做了网络分析,探讨了文章与贡献者之间的互动的网络结构对UGC质量的影响。

综上所述,目前的研究主要还是集中在词条内容维度、网络维度、编辑维度等单维度视角下的评价,多维度视角下的词条质量评价研究还存在欠缺。同时除

了上述学者研究中使用的特征指标外，百度百科的词条属性中还包含了用户的使用行为特征指标：词条使用者的浏览、分享与点赞数值，而目前基于此类指标的百科词条评价与分类研究却非常少见。因此本研究拟解决两个问题：①如何将用户行为指标纳入百度百科词条评价体系中来；②如何从多维度视角对百度百科词条进行综合评价。

3　研究方法

针对提出的两个研究问题，本文参考了国内外学者提出的百科词条评价特征指标，结合 2018 年百度百科最新版本的用户界面的可获取内容，选取了 4 个维度共 12 个量化指标作为研究对象：

（1）内容维度：内容长度指标即摘要长度、正文长度与图片数。内容长度指标代表了词条内容的丰富程度，相较 Blumenstock 使用单词数量作为指标，研究更细化地区分了摘要与正文的长度，同时加上图片数量，使得数据更有说服力；次要内容指标即参考资料数量、标签数量、infobox 数量。三者并不像词条长度与链接指标直接与词条质量相关，而是从侧面反映词条的细致程度。

（2）网络链接维度：出链与入链数量。出链与入链数量作为链接分析中的常用指标，能够体现词条在整个词条网络的联通程度，拥有高出链、入链量的词条，更容易吸引用户从其他词条转入或转出到其它词条，是体现词条被使用概率的重要指标。

（3）用户使用行为维度：浏览量、点赞量、转发量。百科词条质量的高低与否，很大程度上与用户使用行为相关联，高浏览、点赞、转发的词条相比无人问津的词条，词条质量必然也会存在差别。

（4）编辑维度：相较于词条生成时间，词条编辑次数能更好地度量词条的生命周期，体现出高更新频率的高质量词条与长时间无人维护的低质量词条间的差距。

由此，对于要解决的两个研究问题，本文得出了以下解决方案：

（1）如何将用户行为指标纳入百度百科词条评价体系中来。根据百度百科最新版本的用户界面可获取内容，将浏览量、点赞量和转发量作为衡量用户使用行为的三个主要指标，由此单独提出百度百科词条多维度评价指标体系中的用户使用行为维度。

（2）如何从多维度视角对百度百科词条进行综合评价。本文首先基于研究提出的 4 维度 12 指标体系，抓取 2018 年百度百科共 1 500 万余词条的所有内容并统计指标数值；然后基于熵权法计算每个指标的权重系数，度量每个指标对词条质量的影响比重；最后基于灰色关联分析方法对所有百度百科词条进行综合评价并做出排序。

具体使用的方法有：

3.1　指标权重度量方法

为了衡量每个指标对词条综合评价结果的影响力大小，研究选用熵权法计算每个指标的熵权。

熵权法是基于 Shannon 信息论中信息熵[37]概念提出的一种指标权重赋值方法，可以在排除决策者偏好的情况下评估每个参数的权重。熵权法的计算步骤如下[38]：

（1）对现有百度百科词条指标决策矩阵 D（m 个评价指标，n 个词条样本）进行标准化处理，得到标准化矩阵 R：

$$R = (r_{ij})_{m \times n}$$

其中 r_{ij} 是指第 i 个词条的在第 j 个指标上的标准化取值：

$$r_{ij} = \frac{D_{ij} - \min\{D_{ij}\}}{\max\{D_{ij}\} - \min\{D_{ij}\}},\ 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$$

公式（1）

（2）计算所有指标的熵值 H：

$$H_i = -k\sum_{j=1}^{n} f_{ij} m \ln f_{ij},\ i = 1,2,3,\cdots,m \qquad \text{公式（2）}$$

其中：

$$k = \frac{1}{\ln n}$$

$$f_{ij} = \frac{r_{ij}}{\sum_{i=1}^{n} r_{ij}}$$

（3）计算所有指标的熵权 ω：

$$\omega_i = \frac{1 - H_i}{m - \sum_{j=1}^{n} H_j} \qquad \text{公式（3）}$$

式中 $0 \leqslant \omega_i \leqslant 1$，且 $\sum_{i=1}^{m} \omega_i = 1$。

熵权法计算的参数权重越大，参数对评估结果的影响也就越大。

3.2　词条质量评价方法

灰色关联分析（GRA）模型最初由灰色系统理论的创始人邓聚龙教授提出，他于 20 世纪 80 年代提出了这一理论。邓聚龙教授的 GRA 模型在应用中也被称为邓氏灰色关联分析模型或灰色关联分析模型。目前 GRA 模型广泛应用于多目标优化。本研究中的具体评价步骤如下：

（1）根据选定的评价指标，收集并统计实验数据，构建决策矩阵进行标准化处理（同3.1的标准化过程）。

（2）计算每组词条指标与最优值间的关联系数：

$$\zeta_{ij}=\frac{\min_{i=1}^{m}\{\min_{j=1}^{n}(|z_{ij}-z_{0j}|)\}+\eta\times\max_{i=1}^{n}\{\max_{j=1}^{n}(|z_{ij}-z_{0j}|)\}}{|z_{ij}-z_{0j}|+\eta\times\max_{i=1}^{m}\{\max_{j=1}^{n}(|z_{ij}-z_{0j}|)\}}$$

公式（4）

式中 η 为分辨系数，且 $\eta\in(0,1)$。一般来说，η 的取值决定了关联系数之间的差异大小，η 越小，分辨能力越强，本研究 η 取值为0.5。关联系 ξ 则代表了第i个词条在第j个指标上与最优值的关联程度。

（3）根据熵权法得到的权重系数，计算每个词条的关联度，进而作出综合评价。

3.3 数据获取

研究使用数据的抓取流程如下：

（1）初步抓取：通过分析百度百科站点链接结构，发现存在形如 http://baike.baidu.com/view/00000010.htm 的链接地址，通过遍历数字ID，可以得到一系列初始链接。利用Python的urllib2库，可以将页面下载到本地。

（2）数据清洗：首先确定要爬取的百度百科页面的结构元素例如：主标题、副标题、摘要、正文、图片、内链、外链等。同时建立一个名为Page的类，用于描述上述信息，其中标题等用文本变量表示，链接等用列表表示。

（3）数据入库：建立一个MySQL数据库，库中建立一个名为entry的表，将上述页面结构元素用表中的属性表示，使用pymysql库将清洗得到的数据写入数据库。

（4）二次抓取：通过分析发现，初次抓取得到的页面中包含大量形如/item/abc形式的相对链接，链接目标为其他词条页面。

故以初次页面采集得到的内链作为种子URL地址，再次使用步骤1－3中的方法，进行数据抓取和清洗，数据二次抓取时，需要判断URL是否已经被写入数据库，若存在相应记录则不再重复抓取。

4 研究结果及分析

4.1 指标权值

根据全部百科词条数据生成的决策矩阵，应用熵权法求出所有指标的熵权值如表1与图1所示：

表1 百科词条指标权重表

指标序号	指标1	指标2	指标3	指标4	指标5	指标6
指标名称	摘要字数	正文字数	infobox	标签数量	出链	入链
权重	0.022	0.041	0.011	0.018	0.086	0.171
指标序号	指标7	指标8	指标9	指标10	指标11	指标12
指标名称	参考资料	图片数	浏览量	转发量	点赞量	编辑次数
权重	0.052	0.052	0.160	0.185	0.175	0.026

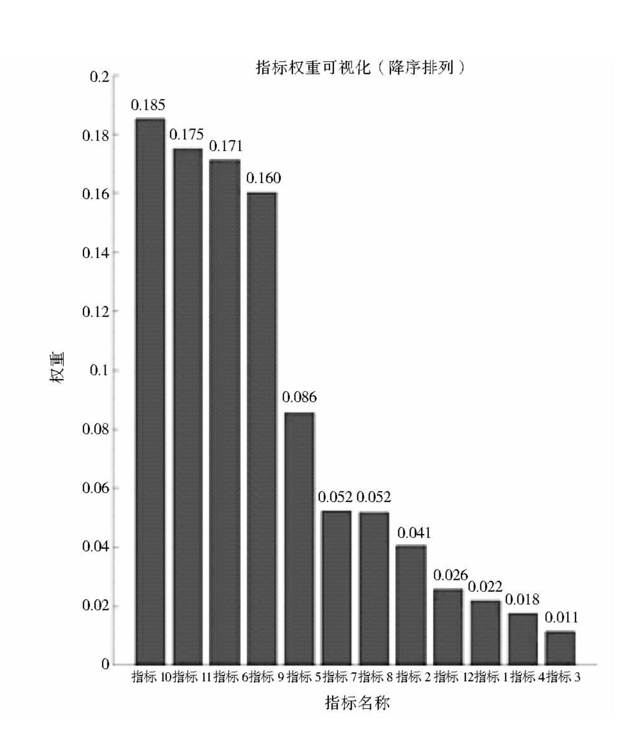


图1 百科词条指标权重

由结果可以看出，权重最高的四个指标包括三个用户使用行为指标——浏览、点赞、转发量，入链数的权重也与三者持平。出链数作为一种“利他”指标，权重不及入链数的一半，但仍然明显高于剩余指标。

4.2　综合评价结果

根据熵权法得出的指标权重，应用到灰色关联分析算法中，继而计算出所有词条的综合评价得分，即词条质量评分，并依此做出排名：

表 2　词条质量排名

词条质量排名前十	台湾（中华人民共和国所属地区），刘诗诗，中央电视台纪录频道，宇智波佐助，南京路（上海南京路），射手座（占星学），后宫动漫，韩国电视剧，DOS（DOS 操作系统），王力宏
词条质量排名后十	相关峰，游戏界面，龙鼓滩发电厂，社会隔离，原生海岸，压缩算法，兼容性测试，汉诺塔问题，单晶硅太阳电池组件，二氧化碳税

高排名词条和低排名词条各维度指标对比如表 3 所示：

表 3　词条指标对比

词条名	摘要长度	正文长度	Infobox 数	标签数	出链数	入链数	参考资料	图片数	浏览量	转发量	点赞量	编辑次数
台湾	779	24 561	30	0	2 213	33 981	242	99	21 386 111	1 661	53 979	3 902
刘诗诗	427	6 522	19	3	347	1 345	131	102	400 725 989	3 287	185 573	3 292
中央电视台纪录频道	172	2 270	9	0	27	2 626	14	5	660 453	93	1 367	2 817
宇智波佐助	394	19 307	26	2	229	454	98	100	12 828 679	815	44 387	2 601
南京路	597	21 395	10	3	384	529	5	9	1 303 856	89	1 796	2 203
相关峰	98	1 060	3	0	3	0	0	8	343	0	0	0
游戏界面	158	1 875	3	1	7	2	0	1	2 438	0	0	0
龙鼓滩发电厂	269	5 015	6	0	1	0	0	1	129	0	0	0
社会隔离	225	978	7	0	4	1	0	1	536	0	0	0
原生海岸	86	645	6	0	2	1	0	1	49	0	0	0

由表 3 可以看出，综合评价排名靠前的词条大部分指标数值都远远高于综合评价评分较低的词条，唯独标签数指标上，各词条并无明显区别，这与熵权法得到的指标权重结果吻合，标签数的权重为 12 个指标中最低。

中央电视台作为综合评分排名第 3 的词条，与前 5 剩余 4 个词条相比，除标签数和编辑次数差异不大，其余指标均不占优势，仅仅靠 2 626 的入链数量便能位居第 3，说明了词条链接网络中优势位置的重要性。

基于熵权法与灰色关联分析做出的词条质量评估，可以清晰体现出不同词条间的质量差别。本研究不仅使用了国内外学者相关研究中使用过的特征指标，还额外加入了用户使用行为指标，结果表明这类指标在词条质量综合评价中影响力与网络链接指标持平，远高于其他指标，因此将用户行为数据纳入百科词条质量评价研究中是有意义的。

5　结语

本研究在结合国内外学者对百科词条研究的理论基础上，基于四个维度提取了百度百科词条的 12 个量化指标，并依此对百度百科词条做出综合评价。

将用户使用行为维度指标纳入词条评价指标体系中是可行且有必要的，三个用户使用行为指标均拥有相对较高的熵权。

（2）用户使用行为指标与网络指标的重要性明显高于其他指标，说明高使用价值与处于词条链接网络优势位置的词条相较于内容优势的词条，影响力更大。

研究所选取的指标都是基于统计的出来的可量化指标，未来将探讨如何将文本丰富度、严谨性、编辑者声誉等定性特征与定量特征相结合。

参考文献：

[1] WIKIPEDIA. Article assessment[EB/OL]. [2018-10-29]. https://en.wikipedia.org/wiki/Wikipedia: Article _ assessment _ (historical).

[2] STVILIA B, FARAJ A, YI Y J. Issues of cross-contextual information quality evaluation -the case of arabic, english, and korean Wikipediaes[J]. Library & information science research, 2009, 31(4):232-239.

[3] LICHTENSTEIN S, PARKER C M. Wikipedia model for collective intelligence: a review of information quality[J]. International journal of knowledge and learning, 2009, 5(3/4):254-272.

[4] ARAZY O, NOV O, PATTERSON R, et al. Information quality in Wikipedia: the effects of group composition and task conflict[J]. Journal of management information systems, 2011, 27(4):71-98.

[5] LIU J, RAM S. Who does what: collaboration patterns in the Wikipedia and their impact on article quality[J]. ACM transactions on management information systems, 2011, 2(2):1-23.

[6] YAARI E, BARUCHSON-ARBIB S, BAR-ILAN J. Information quality assessment of community generated content: a user study of Wikipedia.[J]. Journal of information science, 2011, 37(5):487-498.
[7] OSMAN K. The role of conflict in determining consensus on quality in Wikipedia articles[C]// International symposium on open collaboration. New York: ACM, 2013:1-6.
[8] LEWANDOWSKI D, SPREE U. Ranking of Wikipedia articles in search engines revisited: fair ranking for reasonable quality? [J]. Journal of the American Society for Information Science & Technology, 2014, 62(1):117-132.
[9] 丁敬达.维基百科词条信息质量启发式评价框架研究[J].图书情报知识,2014(2): 11-17.
[10] 赵文轩.网络百科资源质量探析[J].图书馆理论与实践,2016(4): 38-42.
[11] 黄令贺.网络百科信息质量动态演进探索研究[J].情报理论与实践,2018(2):127-131
[12] LIH A. Wikipedia as Participatory journalism: reliable sources? metrics for evaluating collaborative media as a news resource[C]// Proceedings of the 5th international symposium on online journalism. Austin: ISOJ,2004: 16-17.
[13] CHEVALIER F, HUOT S, FEKETE J D. Wikipedia viz: conveying article quality for casual Wikipedia readers[C]// Pacific visualization symposium. Taipei: IEEE, 2010:215-222.
[14] 金燕,周婷,詹丽华.基于层次分析法的协同内容创建系统质量评价体系研究——以百度百科为例[J].图书馆理论与实践,2015(7): 41-45.
[15] 张博,乔欢.协同知识生产社区的内容质量评估模型研究——以维基百科为例[J].现代情报,2015,35(10):17-22.
[16] 金燕.基于用户体验的协同内容创建系统质量保证措施——以百度百科为例[J].情报理论与实践,2016,39(3):6-9.
[17] WARNCKEWANG M, COSLEY D, RIEDL J. Tell me more: an actionable quality model for Wikipedia[C]// International symposium on open collaboration. New York: ACM, 2013:1-10.
[18] STVILIA B, TWIDALE M B, SMITH L C, et al. Assessing information quality of a community-based encyclopedia[C]// International conference on information quality. Cambridge: ICIQ, 2005: 442-454.
[19] BLUMENSTOCK J E. Size matters: word count as a measure of quality on Wikipedia [C]// International conference on World Wide Web. New York: ACM, 2008: 1095-1096.
[20] DALIP D H, CRISTO M. Automatic quality assessment of content created collaboratively by web communities: a case study of Wikipedia[C]// ACM/IEEE-CS joint conference on Digital libraries. New York: ACM, 2009:295-304.
[21] CALZADA G D L, DEKHTYAR A. On measuring the quality of Wikipedia articles [C]// Workshop on Information credibility. New York: ACM, 2010:11-18.
[22] 裘江南,翁楠,徐胜国.基于C4.5的维基百科页面信息质量评价模型研究[J].情报学报, 2012, 31(12):1259-1264.
[23] 李欣奕.网络百科条目质量评价研究[D].长沙:国防科学技术大学, 2014.
[24] 仝召娟,许鑫.百度百科网页质量的自动化评价[J].信息资源管理学报, 2015(2):63-69.
[25] 肖奎,李兵,吴天吉.基于用户行为分析的维基百科词条质量评价方法[J].情报杂志, 2015(5): 185-189.
[26] HU M, LIM E, SUN A, et al. Measuring article quality in Wikipedia: models and evaluation[C]// Conference on information and knowledge management. New York: ACM, 2007:243-252.
[27] ZENG H, ALHOSSAINI M A, DING L, et al. Computing trust from revision history[C]// International conference on privacy, security and trust: bridge the gap between PST technologies and business services. New York: ACM, 2006: 1-1.
[28] LIM E, VUONG B, LAUW H W, et al. Measuring qualities of articles contributed by online communities[C]// IEEE/WIC/ACM International conference on Web intelligence. Washington DC: IEEE, 2006: 81-87.
[29] DONDIO P, BARRETT S. Computational trust in web content quality: a comparative evaluation on the Wikipedia project[J]. Informatica, 2007, 31(2): 151-160.
[30] WOHNER T, PETERS R. Assessing the quality of Wikipedia articles with lifecycle based metrics[C]// International symposium on Wikis. Orlando: DBLP, 2009:1-10.
[31] JAVANMARDI S, LOPES C. Statistical measure of quality in Wikipedia[C]// Proceedings of the first workshop on social media analytics. New York: ACM, 2010:132-138.
[32] YU S, YOSHIKAWA M. Assessing quality score of Wikipedia article using mutual evaluation of editors and texts[C]// ACM International conference on information & knowledge management. New York: ACM, 2013: 1727-1732.
[33] LI X, TANG J, WANG T, et al. Automatically assessing Wikipedia article quality by exploiting article-editor networks[C]// Advances in Information Retrieval, ECIR. Gewerbestrasse: Springer, Cham, 2015:574-580.
[34] WILKINSON D. M, HUBERMAN B. A. Cooperation and quality in Wikipedia[C]// International symposium on Wikis. New York: ACM, 2007: 157-164.
[35] 肖奎,罗保山,吴天吉.一种限定领域的网络文档资源质量评价方法——以维基百科为例[J].情报理论与实践, 2016, 39(12): 120-123.
[36] WU K, ZHU Q, ZHAO Y, et al. Mining the factors affecting the quality of Wikipedia articles[C]// International conference of information science and management engineering. Washington DC: IEEE, 2010:343-346.
[37] 刘思峰,蔡华,杨英杰,等.灰色关联分析模型研究进展[J].系统工程理论与实践, 2013, 33(8):2041-2046.
[38] 倪九派,李萍,魏朝富,等.基于AHP和熵权法赋权的区域土地开发整理潜力评价[J].农业工程学报, 2009, 25(5):202-209.

作者贡献说明:
何振宇:提出研究思路、设计实验方案、数据收集与处理、进行实验、论文撰写;
董祥祥:论文撰写;
朱庆华:提供论文修改意见。

Construction of Baidu Encyclopedia Entry Evaluation Index Under Multi-dimensional Perspective

He Zhenyu　Dong Xiangxiang　Zhu Qinghua
School of Information Management, Nanjing University, Nanjing 210023

Abstract: [**Purpose/significance**] The study incorporates user behavior into the encyclopedia entry evaluation system, and comprehensively evaluates Baidu Encyclopedia entries from a multi-dimensional perspective. [**Method/process**] From four dimensions, the study selected 12 quantitative indicators as research indicators, and comprehensively evaluated Baidu Encyclopedia entries. [**Result/conclusion**] User use behavior indicators and link indicators play a relatively important role in comprehensive evaluation of entries; This study enriches the evaluation system of encyclopedia entry research, but the limitation is that it is difficult to accurately quantify without considering richness and rigor. Characteristics.

Keywords: Baidu Encyclopedia entry　GRA　entropy weight method　comprehensive evaluation